# Final Report

# Learning by Reading for Robust Reasoning in Intelligent Agents

ONR Award #: N00014-16-1-2177

Period of Performance: 2/1/14-01/31/18

ATTN: Dr. Behzad Kamgar-Parsi

**Technical contact**:

Kenneth D. Forbus
EECS Department, Northwestern University
2133 Sheridan Road
Evanston, IL, 60209
Voice: 847-491-7699
Fax: 847-491-5258
Email: forbus@northwestern.edu

**Administrative Contact**:

Andrea Zakrzewski
Associate Director for Grants Management
Office for Sponsored Research
1801 Maple Ave.
2nd Floor, Suite 2410
Evanston, IL 60201-3149
Phone: (847)467-3495
Fax: (847)491-4800
a-zakrzewski@northwestern.edu

# REPORT DOCUMENTATION PAGE

*Form Approved*
*OMB No. 0704-0188*



| **1. REPORT DATE** *(DD-MM-YYYY)* | **2. REPORT TYPE** | **3. DATES COVERED** *(From - To)* |
|---|---|---|
| 04/24/2018 | FINAL TECHNICAL REPORT | 2/1/14-1/31/18 |

**4. TITLE AND SUBTITLE**

Learning by Reading for Robust Reasoning in Intelligent Agents

**5a. CONTRACT NUMBER**

**5b. GRANT NUMBER**

N00014-16-1-2177

**5c. PROGRAM ELEMENT NUMBER**

**6. AUTHOR(S)**

Forbus, Kenneth

**5d. PROJECT NUMBER**

**5e. TASK NUMBER**

**5f. WORK UNIT NUMBER**

**7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)**

Northwestern University, Office for Sponsored Research (OSR)
1801 Maple Ave., 2nd Floor, Suite 2410
Evanston, IL 60201-3149

**8. PERFORMING ORGANIZATION REPORT NUMBER**

**9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)**

Thomas R. Pettit
Office of Naval Research
230 South Dearborn Room 380
Chicago, IL 60604-1595

**10. SPONSOR/MONITOR'S ACRONYM(S)**

ONR

**11. SPONSOR/MONITOR'S REPORT NUMBER(S)**

**12. DISTRIBUTION/AVAILABILITY STATEMENT**

Distribution unrestricted

**13. SUPPLEMENTARY NOTES**

**14. ABSTRACT**

Our hypotheses are that analogical processing plays multiple roles in enabling machines to learn by reading, and that qualitative representations provide a foundation for commonsense knowledge that can be used for reasoning and for natural language semantics. Our accomplishments on this project provide strong evidence for these hypotheses.

**15. SUBJECT TERMS**

Analogy, Qualitative Reasoning, Learning by Reading, Natural Language Understanding

| **16. SECURITY CLASSIFICATION OF:** a. REPORT | b. ABSTRACT | c. THIS PAGE | **17. LIMITATION OF ABSTRACT** | **18. NUMBER OF PAGES** | **19a. NAME OF RESPONSIBLE PERSON** |
|---|---|---|---|---|---|
| None | None | None | UU | 6 | Kenneth D. Forbus |
| | | | | | **19b. TELEPHONE NUMBER** *(Include area code)* 847.491.7699 |

## Goal and Objectives

The goal of this project was to understand how to create intelligent agents that can learn by reading, in order to perform robust commonsense reasoning and interact effectively with human partners. This requires deep understanding of the conceptual knowledge conveyed by texts and constructing knowledge that can be used in subsequent reasoning, by contrast with shallow information-retrieval approaches. It also requires the ability to incrementally adapt to new language usage by their partners, and doing so rapidly, with only a small number of examples (unlike, say, "deep learning" systems that require orders of magnitude more examples than analogical learning systems). Our overall hypotheses are that analogical processing plays multiple roles in learning by reading, and that qualitative representations provide a foundation for common sense knowledge that can be used for reasoning and as part of the semantics for natural language (Forbus et al. 2015). Our accomplishments on this project provides strong evidence for these hypotheses.

## Summary of Accomplishments

### Narrative Function & Qualitative Representations in Semantics

The idea of narrative function is to drive understanding of a text via context, i.e. what should a reader be taking away from reading this text? Narrative function can be seen as a kind of communication act, but the idea goes a bit beyond that. Communication acts are often coarse-grained, e.g. INFORM, QUERY. Narrative functions help bridge between the representations naturally produced via lower levels of language understanding and the rich conceptual representations that need to be inferred from them. Narrative functions provide queries that can be asked of texts. For example, in learning qualitative models from texts, our system begins by looking for occurrences of quantities, then relationships involving them (e.g. changes, ordinal relations, causal influences) and finally processes that connect them. These queries operate abductively, in that they are capable of making assumptions based on the alternatives that are provided by the initial, underspecified semantic representations produced by our parser. The use of abduction for this kind of semantic interpretation is standard, but our contextualized, knowledge-level guidance of the process is novel.

We have shown that our account of narrative function can extract instance-level qualitative representations from texts. In addition to handling the eight examples from our prior work, the system was tested on nine simplified English paragraphs from a science book. On this small corpus, it achieved a precision of 0.74 and recall of 0.45 (McFate et al. 2014). Furthermore, we expanded our account of narrative functions to combine type-level as well as instance-level qualitative representations (McFate et al. 2013, 2014). Instance-level qualitative representations concern specific situations, e.g. "Niagara Falls produces 5.4 gigawatts of electricity". Type-level qualitative representations concern general information, e.g. "Solar panels produce electricity." Flexible understanding of texts requires handling both kinds of knowledge, e.g. instructional texts invariably combine both, to provide both general principles and to tie them into specific examples to help ground them for learners. For our systems, such examples provide base domains for analogies to new situations, as well as inputs to analogical generalization.

We have developed a systematic mapping between constructs of QP theory and FrameNet (McFate & Forbus, 2015), which is interesting for two reasons. First, the ability to create such a mapping provides further evidence that QP theory can play an important inferential role in natural language semantics. Second, since FrameNet provides a very broad coverage of English, tying syntactic constraints on FrameNet frames (called *valence patterns*) to particular QP theory representations sets the stage for radically expanding our coverage of natural language semantics for learning by reading. This analysis also provides, for the first, time, methods of identifying limit points in natural language texts. (In QP theory, a limit point is a threshold where activity changes, e.g. boiling point or pressure equalization.) To test this analysis, we analyzed a new corpus of science articles drawn from the Simple English Wikipedia, and found that 57% of the sentences could be interpreted via QP semantics (McFate & Forbus, 2016b). Over half of those involved lexical units (e.g. "flow") that evoked continuous processes. This provides further evidence for the importance of qualitative representations as a constituent of natural language semantics. These techniques enabled us to successfully extract knowledge from chapter-length texts (McFate & Forbus, 2016c).

## Qualitative Reasoning and Science Tests

Elementary school science tests provide a useful way to measure some aspects of commonsense knowledge. By analyzing a corpus of exams, we found that 29% of such questions involve qualitative reasoning (supporting our hypothesis of its importance in commonsense). In this subset, 13% appeared to be handled by standard qualitative reasoning techniques, and 19% require several kinds of extensions, such as learning specific patterns (e.g. food webs, life cycles). We demonstrated that the standard QR questions can indeed be solved by off-the-shelf techniques, using all 31 questions of that type from the corpus (Crouse & Forbus, 2016).

## Analogical Word Sense Disambiguation

We showed that analogy could be used for both word sense disambiguation and parse tree disambiguation. To record a decision that a specific occurrence of a word (e.g. "star") denoted a specific conceptual sense (e.g. famous human, geometric shape, astronomical body, as drawn from a large off-the-shelf knowledge base), our system automatically constructed cases during natural language understanding that combined the syntactic and semantic information it generated that supported that decision. When understanding subsequent texts, those cases are retrieved and used to provide evidence for interpreting the new text. Our method gets accuracies in the same range as state of the art word sense disambiguation algorithms, but requires very little data. During this project, an expanded version of the conference paper on analogical word sense disambiguation appeared in the journal *Advances in Cognitive Systems* (Barbella & Forbus, 2013).

## Knowledge Segmentation for Analogy-based Learning

How should a system that learns by reading organize what it learns? Obvious possibilities are to store the interpretation of each sentence, paragraph, chapter, or even larger units as a bundle of knowledge. While those surface-level organizations are valuable for readers, they are only indirect markers concerning what aspects of the underlying content are conceptually connected. A topic might be introduced over multiple sentences within a paragraph, with the rest of the paragraph being about something else, for example. Given our hypothesis that analogy provides a basis for human-like robust intelligence, it makes sense to think about how to organize incoming knowledge in terms of cases. We call this the *knowledge segmentation* problem (Barbella & Forbus, 2015). Our approach is to pick out key concepts and/or entities from the understanding of the text, and gather up into a case the facts that mention it (*fact-based segmentation*). Such cases can then be used directly for analogical reasoning (e.g. answering questions about how two things are similar or different) and for analogical learning (e.g. building up a general model of a concept via constructing generalizations based on multiple examples of it). Our fact-based segmentation algorithm out-performed sentence-based algorithms in terms of accuracy on a compare & contrast task based on learned knowledge. Interestingly, a paragraph-based segmentation algorithm produced results that were worse, but not statistically significantly so, from fact-based segmentation. On the other hand, extra facts in the paragraph-based algorithm led to over twice as many analogical matching errors (Barbella & Forbus, 2016). This provides evidence in favor of fact-based segmentation for organizing learned knowledge for subsequent analogical processing.

## Learning Constructions via Analogy

We developed an analogical model for learning the use of denominalized verbs, e.g. "She *spooned* me some sugar." Instead of introducing a new verb, people seem to recognize the intended meaning via a double-object construction, which they associate with transfer events. Given just 21 manually annotated sentences, a Companion was able to learn the double-object construction and then use it to successfully interpret novel verbs that people are able to interpret, as determined by a psycholinguistic study conducted elsewhere (McFate 2016a).

## Other progress on analogy

We published an extended paper on the current version of the Structure-Mapping Engine, our model of analogical mapping and inference, along with the source code and a corpus of over 5,800 comparisons to encourage research in this area (Forbus et al. 2016).

In collaboration with researchers from Google, we showed that analogical matching over relational representations produced by Stanford's Core NLP parser, combined with attributed relational graphs that combine structured information with word2vec statistics, can achieve results competitive with the state of the art in the Microsoft Paraphrase task (Liang et al. 2016). This also illustrates that our approach does not depend on a particular ontology or knowledge base, but is general across relational representations.

Part of our long-range goal, into which this project feeds, is achieving human-level AI via constructing software social organisms. We expanded our formulation by identifying substrate capabilities needed for software social organisms, based on research in other areas of cognitive science (Forbus, 2016).

# Technology Transfer

Our natural-language system has been used by three other projects in our group:

1. ONR Socio-Cognitive Architectures Program: We are using natural language and sketching to communicate simple tasks to Companions, with the goal of being able to instruct them to perform in a variety of psychological experiments, entirely through these natural modalities. We are also exploring learning social reasoning, theory of mind, and moral reasoning through natural language understanding of microstories: Short snippets of text that provide pre-segmented information that might be combined via analogical generalization into useful knowledge.
2. AFOSR Machine Intelligence Program: We developed the idea of type-level qualitative representations in this project, which we are applying more broadly here. That project is also

exploring the use of natural language advice (e.g. "Adding a university to a city increases its science output.") and labeling events during learning by demonstration (e.g. "That was bad." when the trainer's attack failed) as part of exploring how to extend Companions to learn as apprentices working with people in a simulated environment.

3. NSF Spatial Intelligence and Learning Center: We are exploring how to do interactive multimodal knowledge capture, using a combination of simplified English and sketches created with CogSketch (the open-domain sketch understanding system developed in SILC), focusing on analogies commonly used in middle-school science education.

By contrast, our goal in this project was to develop a scientific understanding of how to learn deep conceptual knowledge via reading, and to be able to use that knowledge to reason robustly about the everyday world. We believe that analogical learning provides a more human-like (and hence more scalable and reliable) means of learning from reading than alternatives, such as extracting ground facts via information extraction.

We have three ongoing industrial transitions:

1. We are collaborating with IBM T.J. Watson Research Center, on refining our knowledge base materials into an open-license resource that can benefit the community more broadly, and building a next-generation reasoning system that relies entirely on open-license software.
2. With the help of Forbus and one of his students, Google built their own version of SME for in-house experiments. This led to a joint publication in IJCAI on its use in the Microsoft Paraphrase task.
3. Forbus' group is a Founding Academic Partner in the Platform for Situated Intelligence consortium being created by Microsoft Research. This consortium is providing us with new multimodal dialogue capabilities which we are using in Companions, as well as an opportunity to disseminate our research even more widely.

## Publications


1. Barbella, D. (2016). Analogy in Learning by Reading. Doctoral dissertation, Northwestern University, Department of Electrical Engineering and Computer Science, Evanston, Illinois
2. Barbella, D. and Forbus, K. (2013). Analogical Word Sense Disambiguation. *Advances in Cognitive Systems*, 2:297-315.
3. Barbella, D. and Forbus, K. (2015). Exploiting Connectivity for Case Construction in Learning by Reading. *Proceedings of the Third Annual Conference on Advances in Cognitive Systems*, Atlanta, GA.
4. Barbella, D. and Forbus, K. (2016). Exploiting Connectivity for Case Construction in Learning by Reading. *Advances in Cognitive Systems* 4:169-186
5. Crouse, M., & Forbus, K. (2016) Elementary School Science as a Cognitive System Domain: How Much Qualitative Reasoning is Required? *4th Annual Conference on Advances in Cognitive Systems*, Evanston, IL.
6. Forbus, K. (2016) Software Social Organisms: Implications for Measuring AI Progress. *AI Magazine*, 37(1):85-90.
7. Forbus, K., Barbella, D., McFate, C. (2015). Qualitative Reasoning for Learning by Reading: A Theoretical Analysis. *Proceedings of the 28th International Workshop on Qualitative Reasoning (QR2015)*. Minneapolis, MN.

8. Forbus, K. D., Ferguson, R. W., Lovett, A., & Gentner, D. (2016). Extending SME to handle large-scale cognitive modeling. *Cognitive Science*, DOI: 10.1111/cogs.12377, pp 1-50
9. Liang, C., Paritosh, P., Rajendran, V., and Forbus, K. D. (2016). Learning Paraphrase Identification with Structural Alignment. *Proceedings of IJCAI-2016*.
10. McFate, C.J., Forbus, K.D., and Hinrichs, T.R. (2013). Using Narrative Function to Extract Qualitative Information from Natural Language Texts: A Preliminary Report. *Proceedings of the 27th International Workshop on Qualitative Reasoning*, Bremen, Germany.
11. McFate, C., Forbus, K., & Hinrichs, T. (2014). Using Narrative Function to Extract Qualitative Information from Natural Language Texts. *Proceedings of the 28th AAAI Conference on Artificial Intelligence*, Quebec, Canada.
12. McFate, C.J. and Forbus, K. (2015). Frame Semantics of Continuous Processes. *28th International Workshop on Qualitative Reasoning*, Minneapolis.
13. McFate, C.J. (2015) An Analogical Approach to Learning Constraints on Non-canonical Constructions. *2015 Students of Cognitive Systems Workshop*, Atlanta, GA.
14. McFate, C. & Forbus, K. (2016a). Analogical Generalization and Retrieval for Denominal Verb Interpretation. *Proceedings of the 38th Annual Meeting of the Cognitive Science Society*. Philadelphia, PA.
15. McFate, C. & Forbus, K. (2016b). An Analysis of Frame Semantics of Continuous Processes. *Proceedings of the 38th Annual Meeting of the Cognitive Science Society*. Philadelphia, PA.
16. McFate, C. & Forbus, K. (2016c). Scaling up Linguistic Processing of Qualitative Processes. *Proceedings of the 4th Annual Conference on Advances in Cognitive Systems*. Evanston, IL.